Integrating process safety with molecular modeling-based risk assessment of chemicals within the REACH regulatory framework: Benefits and future challenges

Amanda Lewis, Nikolaos Kazantzis *, Ilie Fishtik, Jennifer Wilcox

Department of Chemical Engineering, Worcester Polytechnic Institute, Worcester, MA 01609-2280, USA

Abstract

Registration, evaluation and authorization of chemicals (REACH) represent a recent regulatory initiative by the European union commission to protect human health and the environment from potentially hazardous chemicals. Under REACH, all stakeholders must submit (thermo)physical, thermochemical, and toxicological data for certain chemicals. The commission’s impact assessment studies estimate that the costs of REACH will be approximately 3–5 billion Euros. The present study advocates the systematic incorporation of computational chemistry and computer-assisted chemical risk assessment methods into REACH to reduce regulatory compliance costs. Currently powerful computer-aided ab initio techniques can be used to generate predictions of key properties of broad classes of chemicals, without resorting to costly experimentation and potentially hazardous testing. These data could be integrated into a centralized IT decision and compliance support system, and stored in a retrievable, easily communicable manner should new regulatory and/or production requirements necessitate the introduction of different uses of chemicals under different conditions. For illustration purposes, ab initio calculations are performed on heterocyclic nitrogen-containing compounds which currently serve as high energy density materials in the chemical industry. Since investigations of these compounds are still in their infancy, stability studies are imperative regarding their safe handling and storage, as well as registration under REACH.

Keywords: Chemicals regulation; Computational chemistry; Chemical process safety; Molecular modeling; Chemical risk

1. Introduction

Registration, evaluation and authorization of chemicals (REACH) form the acronym representing a recent complex regulatory and legislative initiative originally developed and introduced by the European union commission, that aims at protecting human health and the environment from potentially hazardous classes of chemicals. At the same time, REACH aims at stimulating innovation and R&D activity towards the design of safer chemicals and processes, thus enhancing corporate responsibility, as well as promoting competitiveness within the European chemical industry [1–3]. Given the inherent inefficiency and antinomies of the current regulatory framework for chemicals in Europe [1,2], REACH not only represents a comprehensive regulatory policy framework for the management of chemical in the European union (EU), but is also compatible with World Trade Organization (WTO) rules and directives. As a result REACH, will eventually have a much broader impact on chemicals policy and regulation initiatives as they begin to be implemented on a worldwide scale [1–4]. Indeed, REACH policies are going to affect a quite broad group of manufacturers, importers and downstream users of chemical substances [2]. Under the aforementioned regulatory framework, all stakeholders must submit (thermo)physical, thermochemical, toxicological data, as well as the results of risk assessment studies for all chemicals involved through the submission of detailed technical dossiers [2,3,5]. The latter will be thoroughly evaluated by state authorities in all member states of the European union, as well as by the newly established European chemical agency (ECA), and authorization will be issued accordingly for the use and storage of the most hazardous classes of chemicals [2,3,5]. In light of the new legislation and chemicals policy, various impact assessment studies undertaken on behalf of the European commission provide estimates for the associated costs induced by REACH within the range of 3.5 billion Euros [6]. Particular emphasis is placed
on the reduction of the associated regulatory compliance costs within the REACH framework for small to medium-sized enterprises (SMEs) due to their limited resources [2,6]. Taking into account the above considerations, the present study aims at the development of a framework that advocates the systematic incorporation of process safety practices through the use of molecular modeling techniques in order to develop a cost-effective comprehensive computer-assisted chemical risk assessment scheme and integrate it into a centralized supervisory IT-system, the latter being the regulation support system administered by ECA and the European Chemicals Bureau (ECB).

According to the proposed approach, current powerful computer-aided molecular modeling techniques can be used in order to develop and validate quantitative structure-activity relationships (QSARs) [7,8], through which one could computationally generate predictions of key (thermo)physical, thermochemical, and toxicological properties for broad classes of chemicals, as well as assess the associated chemical risks under different conditions without resorting to costly experimentation and potentially hazardous testing. In addition, the computer-based investigations will allow for the reduction of scientifically less sound trial-and-error type of risk assessment and management practices that could induce fines and unnecessary litigation.

The computationally generated data, QSARs and risk assessment results could be integrated into the centralized information management and regulation support system of ECA and ECB, as well as the overall compliance plan and IT-systems of corporations. Preferably, they would be stored in a format that renders the pertinent information retrievable, easily transferable/communicable while facilitating its flow between the various stakeholders should new regulatory and/or production requirements and strategic goals necessitate the introduction of different uses of chemicals under different conditions. Consequently, the preparation of the content of the detailed technical dossiers and compliance to requirements under REACH becomes easier, cost-effective, operationally transparent and amenable to adaptation to new market conditions and regulatory norms. Indeed, preliminary and rather promising results on the cost-saving potential of QSARs under REACH were recently released, further corroborating the intuitive benefits of incorporating process safety and molecular modeling-based risk assessment of chemicals into the new regulatory framework [7–10]. Within the above context and in order to illustrate the proposed approach, molecular modeling investigations based upon quantum mechanics are performed on a heterocyclic nitrogen compound that has recently emerged in the literature due to its promise of serving as a high energy density material (HEDM) in the chemical industry. Since investigations of heterocyclic nitrogen compounds of this type are still in their infancy, stability studies are imperative so that knowledge can be gained regarding their safe handling and storage, as well as their registration under REACH. The present work is the first to examine the formation enthalpy of this novel compound from a theoretical perspective. Future work will involve the examination of other emerging HEDMs in the literature.

The present paper is organized as follows: Section 2 contains a description of the main features, structure and requirements of the new regulatory framework and policy for chemical substances in the EU known as REACH, as well as the main results and findings of recent impact assessment studies on the chemical industry. A few thoughts and ideas on integrating process safety and molecular modeling-based risk assessment of chemicals within REACH, along with the associated benefits and future challenges are presented in Section 3. The proposed ideas are illustrated through a molecular modeling case study in Section 4, followed by some concluding remarks are in Section 5.

2. REACH: a new regulatory and policy framework for chemicals in the European union

It is widely recognized that, the current regulatory framework for the management of chemicals in Europe is inadequate and inefficient [1–3]. In particular, it has not resulted in sufficient information or sound chemical risk assessment practices pertaining to the effects of certain chemicals on human health and the environment. Furthermore, whenever the associated risks of these substances have been identified, the implementation of risk management measures has been unacceptably slow [1–3]. Furthermore, the current framework has adversely affected patterns of research activity and innovation, causing the European chemical industry to lag behind its main counterparts in the US and Japan [1–3].

The currently used regulatory framework makes a clear distinction between the so-called existing and new chemicals. Approximately 100,000 chemicals have been introduced to the global market before 1981 and are termed as existing chemicals, with approximately 3000 being introduced after 1981 and termed as new ones [1,2]. While new chemicals have to undergo extensive testing before entrance into the market, there are no such provisions and comprehensive directives for existing chemicals. The current regulatory framework in the EU requires information on only high volume existing chemicals to be submitted and only public authorities in member states are responsible to determine which of them need further examination [1–3]. As a result, these procedures have been proven to be bureaucratically tedious and inefficient. Current legislation requires manufacturers and importers of chemicals to provide information on the chemicals they use and store, but does not impose similar obligations on downstream users (such as industrial users and formulators) unless the substance is classified [1,2]. Clearly, reliable information on the uses of chemical substances is currently difficult to obtain and information about exposure associated with downstream uses of chemicals is generally scarce. Within the existing framework, new chemicals ought to be notified and tested in production volumes as low as 10 kg/year. This has inhibited R&D activities, undermined invention efforts for new substances, and stifled innovation in the European chemical industry, encouraging the continued use of existing chemicals that current regulation compliance requirements render easier to use and less costly [1,2].

In light of the aforementioned remarks, a revision of the current legislative framework for chemicals in the EU becomes imperative. In response to this need, the EU commission introduced a preliminary White Paper [1], which outlined the main
strategic goals and policy measures for the development of a new regulatory framework for chemicals in Europe. This new ambitious piece of proposed legislation became known under the acronym REACH (registration, evaluation and authorisation of chemicals). Following extensive consultations with major stakeholders, including governments, industry and non-governmental organizations (NGOs), a comprehensive piece of legislation emerged on 29 October 2003 through the commission’s initiatives and put forward for consideration by the European Parliament and Council for possible adoption under the so-called co-decision procedure [2]. The commission’s proposal represents an ambitious model of sustainable development by simultaneously pursuing objectives along three main axes: economic (industrial competitiveness), social (public health protection and job creation), and environmental. The proposal also represents a visible piece of evidence of a growing trend towards increasing corporate responsibility on global regulation requirements, as well as industry-led evaluation and understanding of the risks of chemical exposure and the associated effects on the environment.

At this point, let us present the most salient features of REACH [2]. In the EU, all chemical substances that are manufactured or imported in volumes exceeding one metric tonne on an annual basis per manufacturer or importer (tonnage) must be registered. The registration procedure requires the submission of a technical dossier which contains fundamental information on the chemical’s (thermo)physical, thermochemical, and toxicological properties and uses. It is important to notice that all dossiers will be evaluated and checked. When this procedure is complete, the chemical is considered to be registered and can continue to be used until further evaluation is deemed appropriate. One could single out two special classes of chemical substances that are exempt from current REACH registration requirements for rather obvious reasons: chemical substances solely used and stored for R&D purposes and polymers. Under the proposed legislation, a European chemical agency (ECA) will be established in Helsinki, Finland that will undertake the management of the technical, scientific and administrative aspects of REACH and the data-base of chemical information. The ECA will also ensure that REACH functions well and maintains its credibility and transparency with all stakeholders.

Chemical substances that are manufactured in volumes exceeding 100 metric tonnes per year will be evaluated by state authorities in EU member states and appropriate institutions, who may ask for additional testing and risk assessment studies to be conducted. The newly established ECA will ensure consistency across institutions and state agencies in member states during the evaluation process. The ECA will also provide the requisite IT-capacity and communication protocols for data sharing in order to minimize costs. Furthermore, under REACH, certain chemical substances which are characterized as “substances of very high concern” (carcinogenic mutagenic and toxic to reproduction, persistent bio-accumulative and toxic, persistent organic pollutants) ought to be authorized for specific uses and conditions.

An integral part of the October 2003 REACH proposal pertains to the need of a comprehensive extended impact assessment of the new regulatory framework and the induced cost structure on the competitiveness and innovation capacity of the European chemical industry [6]. Over 40 impact assessment studies have been carried out and made a significant contribution towards a better assessment and understanding of the changes needed in order to achieve a balanced and workable solution for REACH. Let us now briefly examine the main findings that resulted from these studies, starting with the regulatory compliance cost structure. The direct costs induced by REACH are estimated to be within the range of 3-5.2 billion Euros over the first 11 years after the entry into force of the new regulatory framework [6,11]. While the costs induced by the new regulatory framework are certainly real, all impact assessment studies suggest that they are also manageable [6,11]. Further improvement of the testing methods through the development of more efficient practices will result in additional cost reduction. On the other hand, all these studies have also shown that the benefits associated with REACH are substantial [6,11]. In agreement with world bank estimates, these studies indicate that the positive public health and occupational impact of REACH will lead to potential health benefits and savings evaluated at approximately 50 billion Euros over a 30-year period due to the reduced burden associated with various diseases caused by chemicals.

It should be pointed out, that SMEs can be particularly affected by REACH due to their limited financial capacity, resources and weaker market position that can pose major challenges to their regulatory compliance efforts [6]. However, SMEs play a strategically important role in the EU economy and the European chemical industry. In light of this recognition, REACH has already introduced lighter requirements since most SMEs are likely to fall into the category of downstream users. Moreover, SMEs that produce substances are likely to find themselves within the lower tonnage bands, on which lighter regulatory requirements are imposed. Innovative research-oriented SMEs could also take advantage of the exemption scheme for R&D-used chemicals offered by REACH. Finally, the benefits associated with the development of a comprehensive user-friendly IT-support system that will be administered by ECA (and developed in consultation with all stakeholders) will be considerable.

The regulatory compliance cost structure and the aforementioned findings of the various impact assessment studies of REACH provide ample motivation for the development of new approaches. These approaches could improve the cost efficiency of the new regulatory framework while maintaining the overall objectives of REACH. In the present paper, the incorporation of process safety practices and molecular modeling-based risk assessment techniques for chemical substances within REACH is advocated as a potential means to enhance its cost efficiency, functionality, transparency, and most importantly, improve and strengthen the scientific/technical basis of a comprehensive chemicals policy. In the following section, it is argued that the above approach may entail considerable benefits to the adoption and actual implementation of REACH, and at the same time, pose interesting challenges and opportunities for further reflection towards the constant refinement and improvement of the new chemicals policy.
3. Integrating process safety and molecular modeling within REACH: benefits and main challenges

(i) Laboratory tests and experimental studies by resorting to animal testing (in vivo) and/or cell cultures (in vitro).

(ii) The establishment of qualitative structure-activity relationships (SARs) or quantitative structure-activity relationships (QSARs).

In the present study the focus is placed on QSARs and the role of molecular modeling techniques in their establishment and validation. QSARs also have the potential to reduce regulatory compliance costs and animal testing under REACH. For these reasons, let us view QSARs as mathematical representations through which quite complex relationships between intrinsic molecular structural characteristics of a substance and its chemical and biological activity can be modeled [7,9,10]. The intrinsic molecular characteristics that define the structure of a chemical substance play the role of “independent variables” often called molecular descriptors. The data associated with the observed chemical and biological activity/behavior of substances (please see the above classification of different types of data) represent the values of the “dependent variables” of QSARs [7,9,10,14]. It should be pointed out, that the values of descriptors can be obtained either through experimental studies (which are non-trivial and quite often technically impossible) or calculated with the aid of currently available software packages that allow a thorough quantum-mechanical description and insightful molecular modeling of the chemical of interest [7-10,14,15]. Typical examples of molecular descriptors are dipole moment, charge-bond strength, delocalizability index, mid-point potential, highest positive and negative charge, highest and lowest molecular orbitals, etc [9,10]. Using molecular descriptor data for chemical substances and data obtained through direct observation, QSARs can be developed by applying techniques such as regression analysis, neural networks (typically back-propagation modeling methods) and various classification methods [14]. A preliminary QSAR is typically developed on the basis of a training set of data, and later verified using a validation set of data. It should be emphasized that data obtained using computational chemistry and molecular modeling techniques are systematically used for both training and validation purposes when QSARs are developed [9,10,14]. Having developed and appropriately validated QSARs, the benefits engendered by their use are two-fold:

(i) Predictions can be generated about the chemical and biological activity of substances. These can then be adopted for chemical management, risk assessment, classification and labeling purposes, and become naturally integrated into a regulatory framework such as REACH.

(ii) Useful information will be able to be extracted on how facets of chemical and biological activity are affected by specific inherent structural (molecular) characteristics of the substance under consideration.

The above advantages become even more pronounced in the case of untested and poorly characterized chemical substances that need to be registered and carefully managed under REACH. They also apply in cases where new safer substances need to be developed and produced.
The integration of computational chemistry, molecular modeling and QSARs into the overall regulatory framework of REACH. In accordance to article 23 of the proposed regulatory and policy framework of REACH, vertebrate animal testing should be viewed only as a last resort for the attainment of the main registration and evaluation objectives [2]. Recent analysis performed by ECB scientists suggests that approximately 3.9 million additional animal tests could be potentially used in order to comply with REACH regulation requirements if alternative approaches are not pursued [7,8]. As mentioned in Section 2, the pursuit of alternative cost-effective, scientifically sound testing, and risk assessment methods for chemical substances could significantly reduce and control the regulatory compliance cost structure under REACH. Both EU authorities and ECB quickly responded to an initiative and proposal put forward by the Institute for Health and Consumer Protection (IHCP) for the development of intelligent testing strategies (ITS) [16]. ITS will form a new comprehensive framework aiming at making current testing practices cost-effective and less demanding on the number of animal tests needed. This can be attained by promoting an integrated testing scheme that rationally uses a multitude of alternative approaches, where computational chemistry and QSARs will have a prominent role [16]. Emphasis is placed on the need for more coordinated efforts between industry and regulatory authorities on the development, validation and use of QSARs in the spirit promoted by the REACH legislation and the paradigm of increasing corporate responsibility that it advocates [7,8,14]. Besides the potential of significantly reducing the number of animal tests, computational chemistry and QSARs exhibit the potential to rationalize (and quite often expedite) testing, priority setting and risk assessment procedures for chemical substances. This is done by eliminating the need for additional tests under certain conditions and/or providing scientifically supported guidance towards the selection of the appropriate testing methods and risk management measures. Preliminary results of recent studies undertaken by ECB suggest that 1.3–1.9 million test animals could be saved if QSARs are adopted, and substantial cost savings of the order of 1 billion Euros could be achieved through the above ITS scheme [7,8].

The latter figure far exceeds the estimated 10 million Euros cost associated with industry developing its own QSARs and documenting them through the IT-support system [7,8]. One could mention the opportunity for the enhancement of the innovation capacity of the chemical industry in alignment with the special incentives provided by the REACH legislation to design and synthesize new and safer chemicals. This is a task that could significantly be facilitated through computational chemistry techniques and a judicious use of QSARs. These can be proven to be advantageous in cases where certain substances withdrawal and extensive reformulation becomes likely under REACH, and innovation is critical for the introduction of new substances and risk management methods into the market. Studies mentioned in Section 2 suggest that there are additional benefits associated with the use of computational chemistry. Furthermore, certain SMEs can benefit by the use of computational chemistry tools and QSARs, thus reducing costs, eliminating redundant testing, and rationalizing risk management practices under REACH requirements.

The integration of computational chemistry, molecular modeling and QSARs under the REACH framework poses considerable scientific, technical, implementation and legislative challenges. The latter fall beyond the scope of the present paper. The first major challenge pertains to various validation procedures for QSARs developed with the aid of computational chemistry that can be universally accepted by decision-makers and regulatory authorities as reliable and practically useful [7,8,14]. The organization for economic co-operation and development (OECD) made the first attempt to address these challenges [17]. Even though OECD ensured homogeneity of standards and consistency of criteria by explicitly advocating the use of sound scientific practices and methods [17], the above efforts have not yet resulted in a practical, transparent validation framework that would bring the broadest possible consensus amongst policy makers, various QSAR users and regulators [14,18]. The above project should receive immediate priority since QSARs (and the associated computational chemistry tools) could be directly used to support decision-making and regulatory actions in the management of chemicals [12,13,18]. They need to exhibit relative simplicity in generating predictions, and the domain of their validity, their prediction uncertainty and degree of reliability concerning certain classes of chemicals must be reported in an unambiguous manner as well [14,18]. Statistical methods used for the development and validation of QSARs need to become available in order to ensure transparency and allow future refinements and extensions. Critical to the above efforts, would be the recognition that QSARs developed for the prediction of health effects of chemicals substantially differ from the ones used for the prediction of ecological and environmental effects due to the fundamental differences in the nature of the respective endpoints, the associated data as well as the availability of reliable dose- or exposure-response relationships [12,13,18].

A major future challenge related to a cost-effective implementation of the REACH regulatory framework is the development and design of a comprehensive user-friendly IT decision support system. It would require access by both industry and regulatory authorities, and facilitate their respective decision-making processes in a transparent, objective and cost-effective manner. Critical to the above efforts, would be the recognition that QSARs developed for the prediction of health effects of chemicals substantially differ from the ones used for the prediction of ecological and environmental effects due to the fundamental differences in the nature of the respective endpoints, the associated data as well as the availability of reliable dose- or exposure-response relationships [12,13,18].

The latter figure far exceeds the estimated 10 million Euros cost associated with industry developing its own QSARs and documenting them through the IT-support system [7,8]. One could mention the opportunity for the enhancement of the innovation capacity of the chemical industry in alignment with the special incentives provided by the REACH legislation to design and synthesize new and safer chemicals. This is a task that could significantly be facilitated through computational chemistry techniques and a judicious use of QSARs. These can be proven to be advantageous in cases where certain substances withdrawal and extensive reformulation becomes likely under REACH, and innovation is critical for the introduction of new substances and risk management methods into the market. Studies mentioned in Section 2 suggest that there are additional benefits associated with the use of computational chemistry. Furthermore, certain SMEs can benefit by the use of computational chemistry tools and QSARs, thus reducing costs, eliminating redundant testing, and rationalizing risk management practices under REACH requirements.

The latter figure far exceeds the estimated 10 million Euros cost associated with industry developing its own QSARs and documenting them through the IT-support system [7,8]. One could mention the opportunity for the enhancement of the innovation capacity of the chemical industry in alignment with the special incentives provided by the REACH legislation to design and synthesize new and safer chemicals. This is a task that could significantly be facilitated through computational chemistry techniques and a judicious use of QSARs. These can be proven to be advantageous in cases where certain substances withdrawal and extensive reformulation becomes likely under REACH, and innovation is critical for the introduction of new substances and risk management methods into the market. Studies mentioned in Section 2 suggest that there are additional benefits associated with the use of computational chemistry. Furthermore, certain SMEs can benefit by the use of computational chemistry tools and QSARs, thus reducing costs, eliminating redundant testing, and rationalizing risk management practices under REACH requirements.


G3 theory begins with an optimized geometry calculation for the species of interest the second order Möller Plesset perturbation theory, MP2, and then uses this optimized geometry for calculating single-point energies (SPE) at higher levels of theory, e.g. MP4, QCISD(T), and H1F [19]. The optimized geometry calculation was carried out using the MP2(FU) method with the 6-31G(d) basis set. “FU” refers to “full” and insinuates that all of the electrons are included in the electron correlation calculation. Electron correlation becomes important when considering second-row atoms such as carbon and nitrogen [19,27].

The following SPE calculations are performed on the MP2(FU)/6-31G(d) optimized geometry of the heterocyclic C$_{2}$N$_{10}$ compound: MP4(FC)/6-31G(d), MP4(FC)/6-31+G(d), MP4(FC)/6-31G(2df,p), QCISD(T), FC/CQ1G(d), and MP2(full)/G3Large. “FC” refers to “frozen core” and implies that inner-shells are excluded from the electron correlation calculation, making the calculations less time consuming. The G3Large basis set is an extended Pople basis set which includes both polarization and diffuse functions [19]. These energies are presented in Table 1.

Table 1 also lists the three correction factors that are considered in the G3 theory, i.e. spin-orbit (SO) correction, higher level correction (HLC), and zero-point energy (ZPE) correction. Previous studies have shown that molecular SO correction provides no overall improvement in the accuracy of energy calculations [19]. The compound of focus, C$_{2}$N$_{10}$ and all the reference species are molecules making the SO correction negligible. The HLC is calculated using the following equation:

\[-Δ_{n}B(n_{a}−n_{b}) or −Δ_{n}H(D(n_{a}−n_{b}))\] (1)

where $n_{a}$ and $n_{b}$ are the numbers of $\beta$ and $\alpha$ valence electrons, respectively, $\Delta$ the correction for paired electrons in molecules, $B$ the correction for unpaired electrons in molecules, $C$ the correction for the paired electrons in atoms, and $D$ is the correction for unpaired electrons in atoms.

The total G3 energy, $E_{G3}$, is calculated through the evaluation of (2):

\[E_{G3} = E(\text{MP4(FC)/6-31G(d)}) + Δ(+1) + Δ(2df,p) + Δ(QCI) + Δ + Δ(HLC) + \text{ZPE}\]

4. The theoretical prediction of the thermochernical property, formation enthalpy: determining the stability of emerging heterocyclic nitrogen compounds

Ab initio investigations were carried out at the G3 level of theory [19] and the isodesmic approach [20] was employed for the theoretical prediction of the formation enthalpy for the heterocyclic nitrogen compound, 3,6-diazo-1,2,4,5-tetrazine (C$_{2}$N$_{10}$). These thermochemical predictions allow for the development of QSARs from which the stability of these emerging high energy density materials (HEDM) can be determined. All molecular orbital calculations were carried out using Gaussian 98 and Gaussian 03 software packages [21].

G3 theory developed by Curtiss et al. [19], was chosen to calculate the unknown heat of formation of C$_{2}$N$_{10}$. It is an improved version of G2 and is more accurate when calculating heats of formation [19,22]. More specifically, G3 has been successful in prediction heats of formation data for compounds containing a significant number of carbon, nitrogen, and oxygen atoms [19,22]. Since the current work concerns a compound containing 2 carbon atoms and 10 nitrogen atoms, this composite method was a logical choice for maximizing the accuracy of the theoretical predictions. Not only is the G3 theory computationally less expensive than G2, CCSD(T), and QCISD(T) levels of theory, but it also uses considerably less computational time due to the changing basis sets [19,23–26].

<table>
<thead>
<tr>
<th>Reference species</th>
<th>6-31G(d)</th>
<th>6-31G(2df,p)</th>
<th>MP4(FC)/6-31G(d)</th>
<th>ΔCPI</th>
<th>Δ(HLC)</th>
<th>ZPE</th>
<th>$E_{G3}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>NH$_{3}$</td>
<td>−56.2879758</td>
<td>−0.0902997</td>
<td>−0.1294505</td>
<td>−0.0823403</td>
<td>−0.0073612</td>
<td>−0.025344</td>
<td>0.036162</td>
</tr>
<tr>
<td>CH$_{4}$</td>
<td>−231.531749</td>
<td>−0.0140967</td>
<td>−0.0193512</td>
<td>0.0169012</td>
<td>0.0325073</td>
<td>0.09579</td>
<td>0.106636</td>
</tr>
<tr>
<td>C$<em>{6}$H$</em>{12}$N$_{2}$</td>
<td>−247.552126</td>
<td>−0.0114325</td>
<td>−0.1852784</td>
<td>0.0016899</td>
<td>−0.333279</td>
<td>−0.09579</td>
<td>0.094161</td>
</tr>
<tr>
<td>ortho-C$<em>{5}$H$</em>{9}$N$_{2}$</td>
<td>−263.541854</td>
<td>−0.0169012</td>
<td>−0.183451</td>
<td>0.0034666</td>
<td>−0.3309753</td>
<td>−0.09579</td>
<td>0.080891</td>
</tr>
<tr>
<td>meta-C$<em>{5}$H$</em>{9}$N$_{2}$</td>
<td>−263.576812</td>
<td>−0.0173593</td>
<td>−0.1839439</td>
<td>0.0025901</td>
<td>−0.338166</td>
<td>−0.09579</td>
<td>0.081767</td>
</tr>
<tr>
<td>C$<em>{6}$H$</em>{14}$N$_{2}$</td>
<td>−279.695974</td>
<td>−0.0168248</td>
<td>−0.1853779</td>
<td>0.0031899</td>
<td>0.0031899</td>
<td>0.09579</td>
<td>0.069467</td>
</tr>
<tr>
<td>N$<em>{2}$H$</em>{4}$</td>
<td>−110.333922</td>
<td>−0.0080826</td>
<td>−0.08196</td>
<td>−0.0007504</td>
<td>−0.1293291</td>
<td>0.038166</td>
<td>0.029317</td>
</tr>
<tr>
<td>N$<em>{2}$H$</em>{4}$</td>
<td>−111.478453</td>
<td>−0.0191261</td>
<td>−0.1096655</td>
<td>−0.0012444</td>
<td>−0.1365662</td>
<td>−0.044702</td>
<td>0.051904</td>
</tr>
<tr>
<td>CH$<em>{4}$N$</em>{2}$</td>
<td>−94.345203</td>
<td>−0.0079219</td>
<td>−0.0791078</td>
<td>−0.0031048</td>
<td>−0.1213373</td>
<td>−0.038166</td>
<td>0.042294</td>
</tr>
<tr>
<td>N$_{2}$H</td>
<td>−164.370891</td>
<td>−0.0105092</td>
<td>−0.1056388</td>
<td>0.0009507</td>
<td>−0.195183</td>
<td>−0.051088</td>
<td>0.021857</td>
</tr>
<tr>
<td>N$<em>{2}$H$</em>{2}$</td>
<td>−521.834014</td>
<td>−0.103376</td>
<td>−0.1033763</td>
<td>0.8851032</td>
<td>−1.0700452</td>
<td>−0.185794</td>
<td>0.059413</td>
</tr>
</tbody>
</table>

*Due to the computational expense of the SPE calculations for C$_{2}$N$_{10}$ the G3 theory was modified as detailed in the text.*


\[ \Delta(+) = E[\text{MP4(FC)/6-31G(2df,2p)}] - E[\text{MP4(FC)/6-31G(d)}] \]

\[ \Delta(2df, p) = E[\text{MP4(FC)/6-31G(2df, 2p)}] - E[\text{MP4(FC)/6-31G(d)}] \]

\[ \Delta(\text{QCD}) = E[\text{QCISD(T, FC)/6-31G(d)}] - E[\text{MP4(FC)/6-31G(d)}] \]

\[ \Delta = E[\text{MP4(FU)/G3Large}] - E[\text{MP2(FU)/G3Large}] \]

\[ \sum \nu_i \Delta H_{i}^{\text{exp}} + \sum \nu_0 \Delta H_{0}^{\text{exp}} = \sum \nu_i \Delta H_{i}^{\text{MP2}} + \sum \nu_0 \Delta H_{0}^{\text{MP2}} \]

\[ \Delta H_{i}^{\text{exp}} = \frac{1}{\sum \nu_i \Delta H_{i}^{\text{MP2}} + \sum \nu_0 \Delta H_{0}^{\text{MP2}}} \]

\[ P_{i} \pi_{11} \pi_{12} \ldots \pi_{1n} \pi_{01} \ldots \pi_{0q} \]

where

\( \Delta(+) \) is the enthalpy of formation of the reference species.

\( \Delta(2df, p) \) is the enthalpy of formation of the reference species employing the isodesmic approach.

\( \Delta(\text{QCD}) \) is the enthalpy of formation of the reference species employing the QCISD(T) method.

\( \Delta \) is the total enthalpy of formation of the reference species.

\( N \) is the number of species.

\( \nu_i \) and \( \nu_0 \) are the stoichiometric coefficients for species \( i \) and \( 0 \), respectively.

\( \Delta H_{i}^{\text{exp}} \) and \( \Delta H_{0}^{\text{exp}} \) are the experimental enthalpies of formation.

\( \Delta H_{i}^{\text{MP2}} \) and \( \Delta H_{0}^{\text{MP2}} \) are the MP2 predicted enthalpies of formation.

\( P \) is the bond order.

\( \pi \) is the bond order.

\( \nu \) is the stoichiometric coefficient.

\( i, j, k, \ldots \) are the indices for the species.

\( q \) is the number of species.

\( n \) is the number of bonds in the reference species.

\( s \) is the number of species in the reference species.

\( \text{QCISD(T)} \) is the quantum chemical method.

\( \text{MP2(FU)/G3Large} \) is the quantum chemical method.

\( \text{MP4(FC)/6-31G(d)} \) is the quantum chemical method.

\( \text{MP4(FU)/G3Large} \) is the quantum chemical method.

\( \text{QCISD(T, FC)/6-31G(d)} \) is the quantum chemical method.

\( \text{MP4(FC)/6-31G(d)} \) is the quantum chemical method.

\( \text{MP4(FC)/6-31G(d)} \) is the quantum chemical method.

\( \Delta H_{i}^{\text{exp}} \) is the experimental enthalpy of formation of species \( i \).

\( \Delta H_{0}^{\text{exp}} \) is the experimental enthalpy of formation of species 0.

\( \Delta H_{i}^{\text{MP2}} \) is the MP2 predicted enthalpy of formation of species \( i \).

\( \Delta H_{0}^{\text{MP2}} \) is the MP2 predicted enthalpy of formation of species 0.

\( P \) is the bond order.

\( \pi \) is the bond order.

\( \nu \) is the stoichiometric coefficient.

\( i, j, k, \ldots \) are the indices for the species.

\( q \) is the number of species.

\( n \) is the number of bonds in the reference species.

\( s \) is the number of species in the reference species.

\( \text{QCISD(T)} \) is the quantum chemical method.

\( \text{MP2(FU)/G3Large} \) is the quantum chemical method.

\( \text{MP4(FC)/6-31G(d)} \) is the quantum chemical method.

\( \text{MP4(FU)/G3Large} \) is the quantum chemical method.

\( \text{QCISD(T, FC)/6-31G(d)} \) is the quantum chemical method.

\( \text{MP4(FC)/6-31G(d)} \) is the quantum chemical method.

\( \Delta H_{i}^{\text{exp}} \) is the experimental enthalpy of formation of species \( i \).

\( \Delta H_{0}^{\text{exp}} \) is the experimental enthalpy of formation of species 0.

\( \Delta H_{i}^{\text{MP2}} \) is the MP2 predicted enthalpy of formation of species \( i \).

\( \Delta H_{0}^{\text{MP2}} \) is the MP2 predicted enthalpy of formation of species 0.

\( P \) is the bond order.

\( \pi \) is the bond order.

\( \nu \) is the stoichiometric coefficient.

\( i, j, k, \ldots \) are the indices for the species.

\( q \) is the number of species.

\( n \) is the number of bonds in the reference species.

\( s \) is the number of species in the reference species.

\( \text{QCISD(T)} \) is the quantum chemical method.

\( \text{MP2(FU)/G3Large} \) is the quantum chemical method.

\( \text{MP4(FC)/6-31G(d)} \) is the quantum chemical method.

\( \text{MP4(FU)/G3Large} \) is the quantum chemical method.

\( \text{QCISD(T, FC)/6-31G(d)} \) is the quantum chemical method.

\( \text{MP4(FC)/6-31G(d)} \) is the quantum chemical method.

\( \Delta H_{i}^{\text{exp}} \) is the experimental enthalpy of formation of species \( i \).

\( \Delta H_{0}^{\text{exp}} \) is the experimental enthalpy of formation of species 0.

\( \Delta H_{i}^{\text{MP2}} \) is the MP2 predicted enthalpy of formation of species \( i \).

\( \Delta H_{0}^{\text{MP2}} \) is the MP2 predicted enthalpy of formation of species 0.

\( P \) is the bond order.

\( \pi \) is the bond order.

\( \nu \) is the stoichiometric coefficient.

\( i, j, k, \ldots \) are the indices for the species.

\( q \) is the number of species.

\( n \) is the number of bonds in the reference species.

\( s \) is the number of species in the reference species.

\( \text{QCISD(T)} \) is the quantum chemical method.

\( \text{MP2(FU)/G3Large} \) is the quantum chemical method.

\( \text{MP4(FC)/6-31G(d)} \) is the quantum chemical method.

\( \text{MP4(FU)/G3Large} \) is the quantum chemical method.

\( \text{QCISD(T, FC)/6-31G(d)} \) is the quantum chemical method.

\( \text{MP4(FC)/6-31G(d)} \) is the quantum chemical method.

\( \Delta H_{i}^{\text{exp}} \) is the experimental enthalpy of formation of species \( i \).

\( \Delta H_{0}^{\text{exp}} \) is the experimental enthalpy of formation of species 0.

\( \Delta H_{i}^{\text{MP2}} \) is the MP2 predicted enthalpy of formation of species \( i \).

\( \Delta H_{0}^{\text{MP2}} \) is the MP2 predicted enthalpy of formation of species 0.

\( P \) is the bond order.

\( \pi \) is the bond order.

\( \nu \) is the stoichiometric coefficient.

\( i, j, k, \ldots \) are the indices for the species.

\( q \) is the number of species.

\( n \) is the number of bonds in the reference species.

\( s \) is the number of species in the reference species.

\( \text{QCISD(T)} \) is the quantum chemical method.

\( \text{MP2(FU)/G3Large} \) is the quantum chemical method.

\( \text{MP4(FC)/6-31G(d)} \) is the quantum chemical method.

\( \text{MP4(FU)/G3Large} \) is the quantum chemical method.

\( \text{QCISD(T, FC)/6-31G(d)} \) is the quantum chemical method.
The final enthalpy of formation of \( B_0 \) is determined as the average over a complete set of isodesmic RERs.

As an example, consider the evaluation of the ab initio enthalpy of formation of \( C_2N_2H_6 \). The structural formula of this species as well as a set of possible reference species is presented in Fig. 1. As can be seen, \( C_2N_2H_6 \) involves five types of bonds, namely, C–C, C–H, C–N, N–H, N–N, and N=N. The simplest species that involve the last three types of bonds are hydrazine (\( N_2H_4 \)), diazine (\( N_2H_3 \)) and hydrogen azide (\( HN_3 \)). Since these species also involve the bond N–H, it is necessary to add at least one reference species that involve this type of bond, e.g., ammonia (\( NH_3 \)). The only species that involve the bonds C–N and C=N and for which accurate thermochemical data are available are methyamine (\( CH_3N \)), pyridine (\( C_5H_4N \)), pyridazine, 1,3,5-triazine (\( C_3H_3N_3 \)). The last three species involve additionally, C=C and C–H bonds that can be balanced with benzene (\( C_6H_6 \)). Thus, the isodesmic reaction scheme for \( C_2H_2 \) involves 10 reference species and a total of nine types of bonds as shown in Fig. 1.

It is important to note that there have been very few investigations involving \( C_2N_2H_6 \). To the authors’ knowledge this species has not been isolated in the laboratory and, therefore, no experimental data exists for it. In addition, there were limited experimental gas-phase thermochemical data available for the reference species. In particular, the experimental formation enthalpy for \( CH_3N \) has an error bar associated with it of ±8 kcal/mol. Although the current investigation does not examine the effect of the complete error range, it will be considered in future work.

For the compound \( N_2H_4 \), there were multiple experimental formation enthalpies available from the NIST–JANAF thermochemical database [37,38] and the most recently investigated in the literature was used in the calculations for the current work.

The bond matrix generated based on this selection of reference species is presented in Table 2. It may be easily checked that the rank of the bond matrix is equal to 8 and, consequently, only 8 types of bonds from a total of 9 are linearly independent.
Further, a RER involves no more than $8+1=9$ species, one of which should be $C_2N_{10}$. The remaining 8 species may be selected from a total of 10 reference species in $10!/8!/2!=45$ ways, i.e., the total number of isodesmic RERs does not exceed 45 and can be generated using Eq. (5). In reality, due to a specific stoichiometric structure of the system, only four RER out of 45 are stoichiometrically distinct. These are, 

$$3N_2H_4 + 4C(H)_3N_1 + 3(NH_2 + 6HN)$$

$$\rightarrow 6NH_3 + 6CH_4N + 3C_2N_{10} \quad (18)$$

$$3N_2H_4 + 8C(H)N_2 + 3(NH_2 + 6HN)$$

$$\rightarrow 6NH_3 + 6CH_4N + 4C_2H(N) + 3C_2N_{10} \quad (19)$$

$$3N_2H_4 + 6C(H)_2N_2 + 3(NH_2 + 6HN)$$

$$\rightarrow 6NH_3 + 6CH_4N + 2C_2H_6 + 3C_2N_{10} \quad (20)$$

It should be noticed that from two different species with the same brutto-formula $C_2H(N)N_2$ but different structures, i.e., pyridazine and 1,3-diazine, only the second appears in the isodesmic RERs.

Once a complete set of RERs is available, the enthalpy of formation of $C_2N_{10}$ may be readily evaluated using the formalism described above. The necessary experimental gas-phase thermochemical data along with the ab initio-generated gas-phase output data is presented in Table 3. Using these data, the enthalpies of formation of $C_2N_{10}$ obtained from the above four isodesmic RERs are: 739.042, 744.493, 743.444 and 740.296 kcal/mol, respectively, that gives an average value of 741.819 kcal/mol.

### 5. Concluding remarks

Registation, evaluation and authorization of chemicals (REACH) represents a recent regulatory and policy framework for chemicals proposed by the European union commission to protect human health and the environment. The commission’s impact assessment studies estimate that the direct costs of REACH will be of the order of 3–5 billion Euros. In light of the above considerations, a few ideas and thoughts were presented advocating the development of a framework that allows for the systematic incorporation of molecular modeling and computer-assisted risk assessment methods of hazards posed by chemicals into REACH to reduce regulatory compliance costs. According to the proposed approach, currently available and powerful computer-aided molecular modeling techniques can be used to computationally generate predictions of key (thermo)physical, thermochemical, and toxicological properties of wide classes of chemicals, without resorting to costly experimentation and potentially hazardous testing. The above computationally generated data could be integrated into a centralized IT decision and compliance support system. To illustrate the proposed approach, a molecular modeling investigation was presented...
as an example. The investigation involved the theoretical for
mitigation of the efficacy of the models. With the novel het
compound, 3,6-di(diazido)-1,2,4,5-tetrazine (C4N8), that might
have promise as a stable HEDM. Stability calculations involving
nitrogen-containing HEDMs of this type require prior thermo
chemical knowledge, such as formation enthalpies. Due to the
potential instability of these compounds, very few experimental
studies are available. It is quite possible that molecular mod
eling investigations will serve as the bridge to understanding
the behaviour and activity of these types of compounds. This
knowledge can then be applied to methods involving their safe
handling and storage, as well as their registration under REACH.

Acknowledgements

The present research work was presented at the 2005 Pro
cess Safety Symposium, Mary Kay O’Connor Process Safety
Center, Department of Chemical Engineering, Texas A&M Uni
versity, College Station, Texas. The authors would like to
thank the Center’s Director Professor S. Mannan and its staff for the
kind invitation and wonderful hospitality. They would also like
to thank the Boston University Scientific Computing Group for
supercomputer time under a Mariner contract on the IBM pSeries
690 (Regatta) and the IBM pSeries 655 systems. Financial sup
port from the National Science Foundation through grant CTS-
9403432 is gratefully acknowledged by Nikolaos Kazantzis.

References

a Future Chemicals Policy, COM. 88 (2001), CEC, Brussels.
Concerning the Regulation, Evaluation, Authorisation and Restric
tions of Chemicals (2003), CEC, Brussels. (http://europa.eu.int/comm/enviro
ue.int/comm/environment/chemicals/reach.htm).
[3] K. Greiner, J.A. Tichner, New Directions in European Chemicals Policies:
Drivers, Scope and Status, Report (2003) University of Massachusetts at
Lowell, Center for Sustainable Production, Chemicals Policy Initiative,
Lowell, MA (http://www.chemicalspolicy.org).
(2005) 45.
ment, SEC. 117/3 (2003), Working paper, Brussels.
Centre, European Chemicals Bureau, European Commission, Directorate
General, Brussels.
[8] Commission of the European Communities, Extended Impact Assess
ment, SEC. 117/3 (2003), Working paper, Brussels.
[10] S.B. Saraf, W.J. Rogers, D.M. Purn, M.S. Mannan, Integrating molecular
modeling and process safety research, Fluid Phase Equilib. 222-223
A.P. Worth, Use of QSARs in international decision-making frame
works to predict ecologic effects and environmental fate of chemical
A.P. Worth, Use of QSARs in international decision-making frameworks
to predict health effects of chemical substances, Environ. Health Pro
[14] L. Erickson, J. Jaworska, A.P. Worth, M.D. Cronn, R.M. McDowell,
P. Gramatica, Methods for reliability and uncertainty assessment and for
applicability evaluations of classification- and regression-based QSARs,
[16] European Chemicals Bureau, Newsletter, Institute for Health and Con
Chemicals Program, in The Use of QSARs for Chemicals Screening,
Inspectorate, Stockholm, Sweden.
workshop on regulatory acceptance of (Q)SARs for human health and
[19] L.A. Curtius, K. Raghavaraju, P. Rehfellin, V. Russelons, J.A. Pople,
Gaussian-3 (G3) theory for molecules containing first- and second-row
[20] W.J. Hehe, R. Ditchfield, L. Radom, J.A. Pople, Molecular orbital the
ory of the electronic structure of organic compounds. IV. Internal rotation
in hydrocarbons using a minimal Slater-type basis, J. Am. Chem. Soc.
[21] M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb,
J.R. Cheeseman, J.A. Montgomery Jr., T. Vreven, K.N. Kudin, J.C.
Burant, J.M. Maren, S.S. Joosens, J. Tomasi, V. Barone, B. Mennucci,
M. Cossi, G. Scalmani, N. Rega, G.A. Petersson, H. Nakatsuji, M.
Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T.
Nakajima, Y. Honda, O. Kitao, H. Nakajima, M. Klene, X. Li, J.E. Kastr,
H.P. Hratchian, J.B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Com-
pers, R.E. Stratmann, O. Yazyev, A.J. Austin, R. Cammi, C. Pomelli,
Dannenberg, V.G. Zakrzewski, S. Dapprich, A.D. Daniels, M.C. Span,
O. Furcik, D.K. Mckie, D.D. Rabuck, K. Raghavaraju, J.B. Foresman,
J.V. Ortiz, Q. Cui, A.G. Baboul, S. Clifford, J. Ciosdowski, B.B. Ste
fanov, G. Liu, L. Liashenko, P. Piskorski, I. Komaromi, R.L. Martin, D.J.
Fox, T. Keith, M.A. Al-Laham, C.Y. Peng, A. Nanayakkara, M. Challis
combe, P.M.W. Gill, B. Johnson, W. Chen, M.W. Wong, C. Gonzalez,
J.A. Pople, Gaussian-3, Revision C02, Gaussian, Inc. Wallingford, CT,
2004.
[22] S. Hammarner, Heats of formation and proton affinities by the G3
[23] L.A. Curtius, K. Raghavaraju, J.A. Pople, Gaussian-2 theory using
[24] B.J. Smith, L. Radom, Calculation of proton affinities using the G2/MP2,
[28] R. Krishnan, J.A. Pople, Approximate fourth-order perturbation theory
[29] L.A. Curtius, P. Rehfellin, D.J. Pumpl, Theoretical methods for com
puting enthalpies of formation of gaseous compounds, Rev. Comput.
effects on the stability of protonated benzene, J. Am. Chem. Soc. 96


